STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _/0/563,976_
Source: _1 Fwp_
Date Processed by STIC: _1/23/06_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT
MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (**<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , EFS Submission **User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):**
   U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

Raw Sequence Listing Error Summary

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics The number/text at the end of each line "wrapped" down to the next line. This may occur if your file
   Wrapped Aminos was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will
   prevent "wrapping."

2 ____ Invalid Line Length The rules require that a line **not exceed** 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino The numbering under each 5th amino acid is misaligned. Do **not** use tab codes between numbers;
   Numbering use **space characters**, instead.

4 ____ Non-ASCII The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please
   ensure your subsequent submission is saved in ASCII text.**

5 ____ Variable Length Sequence(s) ____ contain n's or Xaa's representing more than one residue. **Per Sequence Rules,
   each n or Xaa can only represent a single residue.** Please present the **maximum** number of each
   residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0 A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
   "bug" sequences(s) _____ . Normally, PatentIn would automatically generate this section from the
   previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to
   the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for
   Artificial or Unknown sequences.**

7 ____ Skipped Sequences Sequence(s) _____ missing. If intentional, please insert the following lines for **each** skipped sequence:
   (OLD RULES) (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
   (i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
   (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
   This sequence is intentionally skipped

   Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8 ____ Skipped Sequences Sequence(s) _____ missing. If **intentional**, please insert the following lines for **each** skipped sequence.
   (NEW RULES) <210> sequence id number
   <400> sequence id number
   000

9 ____ Use of n's or Xaa's Use of n's and/or Xaa's have been detected in the Sequence Listing.
   (NEW RULES) Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
   In <220> to <223> section, please explain location of **n** or **Xaa**, and which residue **n** or **Xaa** represents.

10 ___ Invalid <213> Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or
   Response scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or
   is Artificial Sequence

11 ____ Use of <220> Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
   Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or
   "Unknown." Please explain source of genetic material in <220> to <223> section.
   (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0 Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file,
   "bug" resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence
   listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n/Xaa "n" can **only** represent a single <u>nucleotide</u>; "Xaa" can **only** represent a single <u>amino acid</u>
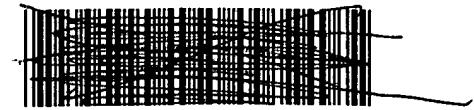
IFWP

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/563,976

DATE: 01/23/2006
TIME: 09:57:37

Input Set : **A:\HOI-14402.ST25.txt**
Output Set: **N:\CRF4\01232006\J563976.raw**

```
     3 <110> APPLICANT: Sorensen, Anders Per
     4         Benfield, Thomas Lars
     5         Lundgren, Jens Dilling
     6         Kempe, Thomas D.
     8 <120> TITLE OF INVENTION: BINDING MEMBER TOWARDS PNEUMOCOCCUS SURFACE
ADHESIN A PROTEIN
     9         (PsaA)
    11 <130> FILE REFERENCE: HOI-14402/16
C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/563,976
C--> 13 <141> CURRENT FILING DATE: 2006-01-09
    13 <150> PRIOR APPLICATION NUMBER: PCT/DK04/000492
    14 <151> PRIOR FILING DATE: 2004-07-08
    16 <150> PRIOR APPLICATION NUMBER: US 60/486,647
    17 <151> PRIOR FILING DATE: 2003-07-11
    19 <150> PRIOR APPLICATION NUMBER: PA 2003 01044
    20 <151> PRIOR FILING DATE: 2003-07-08
    22 <160> NUMBER OF SEQ ID NOS: 56
    24 <170> SOFTWARE: PatentIn version 3.3
    26 <210> SEQ ID NO: 1
    27 <211> LENGTH: 33
    28 <212> TYPE: DNA
    29 <213> ORGANISM: Homo sapiens
    32 <220> FEATURE:
    33 <221> NAME/KEY: CDS
    34 <222> LOCATION: (1)..(33)
    35 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
    37 <400> SEQUENCE: 1
    38 cgg gcg agt cag ggt att agc agc tgg tta gcc                        33
    39 Arg Ala Ser Gln Gly Ile Ser Ser Trp Leu Ala
    40 1               5                   10
    43 <210> SEQ ID NO: 2
    44 <211> LENGTH: 11
    45 <212> TYPE: PRT
    46 <213> ORGANISM: Homo sapiens
    48 <400> SEQUENCE: 2
    50 Arg Ala Ser Gln Gly Ile Ser Ser Trp Leu Ala
    51 1               5                   10
    54 <210> SEQ ID NO: 3
    55 <211> LENGTH: 21
    56 <212> TYPE: DNA
    57 <213> ORGANISM: Homo sapiens
    60 <220> FEATURE:
    61 <221> NAME/KEY: CDS
    62 <222> LOCATION: (1)..(21)
```

Does Not Comply
Corrected Diskette Needed

```
 63 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
 65 <400> SEQUENCE: 3
 66 gtt gca tcc agt ttg caa agt                                          21
 67 Val Ala Ser Ser Leu Gln Ser
 68 1               5
 71 <210> SEQ ID NO: 4
 72 <211> LENGTH: 7
 73 <212> TYPE: PRT
 74 <213> ORGANISM: Homo sapiens
 76 <400> SEQUENCE: 4
 78 Val Ala Ser Ser Leu Gln Ser
 79 1               5
 82 <210> SEQ ID NO: 5
 83 <211> LENGTH: 27
 84 <212> TYPE: DNA
 85 <213> ORGANISM: Homo sapiens
 88 <220> FEATURE:
 89 <221> NAME/KEY: CDS
 90 <222> LOCATION: (1)..(27)
 91 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
 93 <400> SEQUENCE: 5
 94 caa cag tat aat agc tat cct ccg acg                                  27
 95 Gln Gln Tyr Asn Ser Tyr Pro Pro Thr
 96 1               5
 99 <210> SEQ ID NO: 6
100 <211> LENGTH: 9
101 <212> TYPE: PRT
102 <213> ORGANISM: Homo sapiens
104 <400> SEQUENCE: 6
106 Gln Gln Tyr Asn Ser Tyr Pro Pro Thr
107 1               5
110 <210> SEQ ID NO: 7
111 <211> LENGTH: 321
112 <212> TYPE: DNA
113 <213> ORGANISM: Homo sapiens
116 <220> FEATURE:
117 <221> NAME/KEY: CDS
118 <222> LOCATION: (1)..(321)
119 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
V-segment: 4-34
120        and J-segment: JK1
122 <220> FEATURE:
W--> 123 <221> NAME/KEY: CDR1
124 <222> LOCATION: (70)..(120)
126 <220> FEATURE:
W--> 127 <221> NAME/KEY: CDR2
128 <222> LOCATION: (148)..(168)
130 <220> FEATURE:
W--> 131 <221> NAME/KEY: CDR3
132 <222> LOCATION: (265)..(291)
```

RAW SEQUENCE LISTING   DATE: 01/23/2006
PATENT APPLICATION: US/10/563,976   TIME: 09:57:37

Input Set : A:\HOI-14402.ST25.txt
Output Set: N:\CRF4\01232006\J563976.raw

```
134 <400> SEQUENCE: 7
135 gac atc cag atg acc cag tct cca tcc tca ctg tct gca tct gta gga      48
136 Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
137 1               5                   10                  15
139 gac aga gtc acc atc act tgt cgg gcg agt cag ggt att agc agc tgg      96
140 Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Trp
141             20                  25                  30
143 tta gcc tgg tat cag cag aaa cca gag aaa gcc cct gag tcc ctg atc     144
144 Leu Ala Trp Tyr Gln Gln Lys Pro Glu Lys Ala Pro Glu Ser Leu Ile
145         35                  40                  45
147 tat gtt gca tcc agt ttg caa agt ggg gtc cca tca agg ttc agc ggc     192
148 Tyr Val Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
149     50                  55                  60
151 agt gga tct ggg aca gat ttc act ctc acc atc agc agc ctg cag cct     240
152 Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
153 65                  70                  75                  80
155 gaa gat ttt gca act tat tac tgc caa cag tat aat agc tat cct ccg     288
156 Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ser Tyr Pro Pro
157                 85                  90                  95
159 acg ttc ggc caa ggg acc aag gtg gaa atc aaa                         321
160 Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
161             100                 105
164 <210> SEQ ID NO: 8
165 <211> LENGTH: 107
166 <212> TYPE: PRT
167 <213> ORGANISM: Homo sapiens
169 <400> SEQUENCE: 8
171 Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
172 1               5                   10                  15
175 Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Trp
176             20                  25                  30
179 Leu Ala Trp Tyr Gln Gln Lys Pro Glu Lys Ala Pro Glu Ser Leu Ile
180         35                  40                  45
183 Tyr Val Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
184     50                  55                  60
187 Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
188 65                  70                  75                  80
191 Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ser Tyr Pro Pro
192                 85                  90                  95
195 Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
196             100                 105
199 <210> SEQ ID NO: 9
200 <211> LENGTH: 15
201 <212> TYPE: DNA
202 <213> ORGANISM: Homo sapiens
205 <220> FEATURE:
206 <221> NAME/KEY: CDS
207 <222> LOCATION: (1)..(15)
208 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
```

```
210 <400> SEQUENCE: 9
211 ggt ttc tcc tgg agc                                              15
212 Gly Phe Ser Trp Ser
213 1             5
216 <210> SEQ ID NO: 10
217 <211> LENGTH: 5
218 <212> TYPE: PRT
219 <213> ORGANISM: Homo sapiens
221 <400> SEQUENCE: 10
223 Gly Phe Ser Trp Ser
224 1             5
227 <210> SEQ ID NO: 11
228 <211> LENGTH: 51
229 <212> TYPE: DNA
230 <213> ORGANISM: Homo sapiens
233 <220> FEATURE:
234 <221> NAME/KEY: CDS
235 <222> LOCATION: (1)..(51)
236 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
238 <400> SEQUENCE: 11
239 gaa atc gat tat aga gga agc acc aac tac aac ccg tcc ctc aag agt      48
240 Glu Ile Asp Tyr Arg Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser
241 1             5                   10                  15
243 cga                                                              51
244 Arg
248 <210> SEQ ID NO: 12
249 <211> LENGTH: 17
250 <212> TYPE: PRT
251 <213> ORGANISM: Homo sapiens
253 <400> SEQUENCE: 12
255 Glu Ile Asp Tyr Arg Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser
256 1             5                   10                  15
259 Arg
263 <210> SEQ ID NO: 13
264 <211> LENGTH: 21
265 <212> TYPE: DNA
266 <213> ORGANISM: Homo sapiens
269 <220> FEATURE:
270 <221> NAME/KEY: CDS
271 <222> LOCATION: (1)..(21)
272 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
274 <400> SEQUENCE: 13
275 ggg ggg ccc cgc ttt gac tac                                      21
276 Gly Gly Pro Arg Phe Asp Tyr
277 1             5
280 <210> SEQ ID NO: 14
281 <211> LENGTH: 7
282 <212> TYPE: PRT
283 <213> ORGANISM: Homo sapiens
```

```
285 <400> SEQUENCE: 14
287 Gly Gly Pro Arg Phe Asp Tyr
288 1               5
291 <210> SEQ ID NO: 15
292 <211> LENGTH: 345
293 <212> TYPE: DNA
294 <213> ORGANISM: Homo sapiens
297 <220> FEATURE:
298 <221> NAME/KEY: CDS
299 <222> LOCATION: (1)..(345)
300 <223> OTHER INFORMATION: Sequence from human antibody generated in mouse.
301       V-segment: 4-34, D-segment: unknown, J-segment: JH4b
303 <220> FEATURE:
W--> 304 <221> NAME/KEY: CDR1
305 <222> LOCATION: (91)..(102)
307 <220> FEATURE:
W--> 308 <221> NAME/KEY: CDR2
309 <222> LOCATION: (148)..(199)
311 <220> FEATURE:
W--> 312 <221> NAME/KEY: CDR3
313 <222> LOCATION: (191)..(312)
315 <400> SEQUENCE: 15
316 cag gtg cga cta cag cag tgg ggc gca gga ctg ttg aag cct tcg gag    48
317 Gln Val Arg Leu Gln Gln Trp Gly Ala Gly Leu Leu Lys Pro Ser Glu
318 1               5                   10                  15
320 acc ctg tcc ctc acc tgc gct gtc ttt ggt ggg tcc ttc agt ggt ttc    96
321 Thr Leu Ser Leu Thr Cys Ala Val Phe Gly Gly Ser Phe Ser Gly Phe
322             20                  25                  30
324 tcc tgg agc tgg atc cgc cag acc cca ggg aag ggg ctg gag tgg atc   144
325 Ser Trp Ser Trp Ile Arg Gln Thr Pro Gly Lys Gly Leu Glu Trp Ile
326             35                  40                  45
328 ggg gaa atc gat tat aga gga agc acc aac tac aac ccg tcc ctc aag   192
329 Gly Glu Ile Asp Tyr Arg Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys
330     50                  55                  60
332 agt cga gtc acc ata tta aga gac acg tcc agg agc cag ttc tcc ctg   240
333 Ser Arg Val Thr Ile Leu Arg Asp Thr Ser Arg Ser Gln Phe Ser Leu
334 65                  70                  75                  80
336 aag ttg agc tcc gtg acc gcc gcg gac tcg gct gtg ttt tat tgt gcg   288
337 Lys Leu Ser Ser Val Thr Ala Ala Asp Ser Ala Val Phe Tyr Cys Ala
338             85                  90                  95
340 aga ggg ggg ccc cgc ttt gac tac tgg ggc cag gga acc ctg gtc acc   336
341 Arg Gly Gly Pro Arg Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr
342             100                 105                 110
344 gtc tcc tca                                                        345
345 Val Ser Ser
346             115
349 <210> SEQ ID NO: 16
350 <211> LENGTH: 115
351 <212> TYPE: PRT
```

```
<210>   17
<211>   33
<212>   DNA
<213>  (synthetic)
```

invalid response. See item 10 on Error Summary Sheet.

This error appears in subsequent sequences, too.

```
<210>  18
<211>  11
<212>  PRT
<213>  synthetic        error

<220>
<221>  misc_feature
<222>  (33)..(33)
<223>  unknown nucleotide

<400>  18
```

delete this — it applies to the previous (DNA) sequence

Arg Ala Ser Gln Ser Val Ser Ser Tyr Leu Ala
1   5   10

same error in Seq. 24

FYI

<u>**Please Note:**</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the
<220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:17; N Pos. 33
Seq#:23; N Pos. 102

Input Set : **A:\HOI-14402.ST25.txt**
Output Set: **N:\CRF4\01232006\J563976.raw**

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:123 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
L:127 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
L:131 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
L:304 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15
L:308 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15
L:312 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15
L:404 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
L:494 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:23
L:503 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:23
L:507 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:23
L:519 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:96
L:692 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:31
L:696 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:31
L:700 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:31
L:873 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:39
L:877 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:39
L:881 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:39
L:1054 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:47
L:1058 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:47
L:1062 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:47